

Viewpoint

Challenges of qualitative data sharing in social sciences

Tanja Vuckovic Juros✉

Faculty of Humanities and Social Sciences at the University of Zagreb, Croatia

tvuckovi@m.ffzg.hr

orcid.org/0000-0001-6186-883X

Received: 8 Nov 2021

Accepted: 14 Feb 2022

Published: 4 Apr 2022

Competing interests

The author declares no competing interests.

Funding statement

No funding was received for this study.



This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0).

Citation

Vuckovic Juros T. Challenges of qualitative data sharing in social sciences. *Eur Sci Ed.* 2022;48:e77781. <https://doi.org/10.3897/ese.2022.e77781>

Abstract

Open science offers hope for new accountability and transparency in social sciences. Nevertheless, it still fails to fully consider the complexities of qualitative research, as exemplified by a reflection on sensitive qualitative data sharing. As a result, the developing patterns of rewards and sanctions promoting open science raise concern that quantitative research, whose “replication crisis” brought the open science movement to life, will benefit from “good science” re-evaluations at the expense of other research epistemologies, despite the necessity to define accountability and transparency in social sciences more widely and not to conflate those with either reproducibility or data sharing.

Keywords:

Accountability, open science, qualitative research, sensitive data, transparency

Introduction

The goal of open science (OS) is to make all research elements, from study inception to final outputs, available and accessible. OS practices include data sharing, open (transparent) methodology, and open access to publications. This is not an exhaustive list (see also open peer review, citizen science, etc.),^{1,2} but these are the practices most highlighted by the funding mandates of various institutions and agencies. Data sharing and open methodology, in particular, have a growing role in social sciences as researchers are pushed to make their data findable, accessible, interoperable, and reusable (FAIR),³ to document their analytic and interpretative pipelines, and to provide access to their research materials. Laudable as they are, these practices are not without their challenges, especially for qualitative research.

Qualitative social scientists have many reasons to support open science, especially data sharing and open methodology. Notably, these practices bring a quantitative analyst back into the picture by shredding the illusion of the variables and numbers' "objectivity" and "neutrality" that quantitative researchers previously used to legitimize their work. The "replication crisis" – that led to the rise of the OS movement – vividly demonstrated how focusing solely on the results and obfuscating the research process can lead to bad science.⁴ Indeed, quantitative analysts are also part of the research process as they make decisions affected by ethics, integrity, and moral values, in addition to issues of competence or carelessness.^{4,5} Therefore, by making quantitative researchers visible in the research process, the OS movement has also put them on a more level ground with qualitative researchers, who have long challenged the concepts of "objectivity" and "neutrality" by demanding reflexivity and positionality in qualitative research.

These demands for reflexivity and positionality stem from the qualitative researchers' preoccupation with transparency and accountability – the understanding that the conclusions are only as good as the data and that the data are only as good as how clearly the analysts can trace their research decisions, analyses, and interpretations (see "audit trail").⁶ Therefore, qualitative research principles are clearly compatible with OS principles.⁷ The developing standards of OS also benefit qualitative research in other ways. For example, journals' word limits often prevent qualitative researchers from fully presenting their analytic and interpretative pipelines. The option to deposit such materials into data repositories thus provides a welcome solution to a long-standing problem. Moreover, qualitative data are notoriously difficult and time-consuming to collect. Having an opportunity to access existing qualitative data sets, for both teaching and new analyses, could profoundly advance qualitative social sciences.

The peril of delegitimizing qualitative research

This enthusiasm notwithstanding, current OS practices also trouble many qualitative researchers.⁷ First, the data sharing requirements are rarely considered with regard to their full implications for research ethics and qualitative data integrity. Further, the open (transparent) methodology is often conflated with reproducible methodology. As a result, the developing pattern of OS rewards and sanctions creates a concern that quantitative social sciences will benefit from new "good science" re-evaluations at the expense of other types of research, including those that preached research transparency and accountability long before the OS movement.

Consider, for example, the perfunctory nods to the qualitative data complexities in the OS phrases such as "as open as possible, as closed as necessary"⁸ or in the guidelines for Transparency and Openness Promotion

(TOP) that suggest that the journals implementing OS can allow for exceptions to data transparency “in rare cases, [when] despite authors’ best efforts, some or all of data or materials cannot be shared for legal or ethical reasons.”⁹ This OS framework of exceptions to the “best” practice does not recognize other types of good practices developed for diverse research epistemologies. Instead, it formally rewards solely the good practices of quantitative research, while the qualitative researchers are disproportionately burdened with justifying their research design – with, for example, defending “as closed as necessary” even when anything else would be inappropriate. In this framework of exceptions – increasingly used by funding agencies and reinforced by academic publishing (e.g., journals can bestow “badges” to articles implementing the TOP guidelines)¹⁰ – qualitative researchers are clearly disadvantaged. There are no visible benefits for recognizing the complexities of qualitative data, neither in funding schemes nor in publishing. In these circumstances, it is easy to imagine that qualitative research (especially innovative and non-conventional research) will increasingly seem a risky endeavor, much to the detriment of social sciences.

A cautionary tale of ethics committees

This concern is grounded in previous experience with the research ethics committees. Developed initially in response to the misdeeds of biomedical research and experimental psychology, formal ethics reviews often imposed the ethical concerns of such research to those with different types of ethical considerations.^{11, 12} Despite the progress in particular academic contexts,¹³ qualitative researchers writing about ethics demonstrate that institutionalized ethics requirements still ask the wrong questions, do not anticipate real problems, and offer no meaningful framework to help qualitative researchers deal with specific ethical issues.^{12, 14} What such

ethics requirements rather do is discourage transformative or challenging qualitative research and create unnecessary obstacles. Qualitative researchers, therefore, have to write their proposals so as to satisfy the ethics committees, rather than address their ethical concerns, and qualitative researchers seldom benefit from the process of ethics evaluation.^{11, 15} This translates into academic publishing as well, as journals follow the cue of institutionalized ethics reviews and rarely expect or offer sufficient space for qualitative researchers to fully address how research ethics was integrated into a qualitative study design beyond the issue of informed consent or removing identifiers.

Transparency vs. reproducibility

A similar problem of inappropriate evaluation criteria for qualitative research is also evident in the OS movement’s focus on reproducibility instead of transparency. While reproducibility has no place in the epistemology of qualitative research, transparency – about assumptions, procedures, and ethical concerns informing analyses and interpretations – is central. But, in the OS movement, developing as it did from the replication crisis, transparency serves reproducibility. This is evident from the opening lines of the TOP guidelines: “Reproducibility of research can be improved by increasing transparency of the research process and products.”⁹ When such assumptions are imposed as the main standard of “good science” and then institutionalized into academic publishing, they disadvantage qualitative researchers. Reproducibility, therefore, should be de-emphasized from the OS vocabulary, lest “open science” become another means of delegitimizing qualitative research, the same purpose that “neutrality” and “objectivity” served previously. Instead, the OS movement must move beyond the language of “exceptions” to fully include qualitative

social sciences by considering how best to address the complexities of qualitative data. This, at the least, includes the understanding that transparency, not reproducibility, is the desired outcome and that transparency does not necessarily require data sharing.

Challenges of sensitive qualitative data sharing

I will illustrate the problem with qualitative data sharing by a short reflection from a perspective of a researcher working with sensitive data. Let us suppose that I am depositing my data into the Qualitative Data Repository (QDR; data.qdr.syr.edu), a well-known US qualitative research archive. The QDR puts great emphasis on the de-identification of qualitative data, where direct identifiers are to be removed and indirect identifiers aggregated, generalized, or otherwise manipulated to protect the confidentiality of research participants.¹⁶ Likewise, the QDR anticipates that the depositors choose between several categories of access, including depositor-approved access for registered QDR users and restricted offline access, both with an embargo option.¹⁷ These are currently the best practices of qualitative data sharing, allowing depositors to preserve a lot of control over their sensitive data. Still, even within this framework, I can highlight several challenges for research ethics and data integrity.

In my example, qualitative data are sensitive personally (the participants share intimate matters) and politically (the topic is at the center of divisive political struggles). Such data are produced in the relationship of trust with the research participants. The participants trust the researchers not to misuse, misrepresent, or misinterpret the data. They also trust that the researchers will respect their ongoing¹⁴ responsibility to the participants. This includes preserving participants' confidentiality in relation to both external observers and those who need fewer cues for recognition.¹² The solution may seem straightforward – de-identification of both the deposited data and

the publications. However, while the data are more easily deidentified in publications without jeopardizing the integrity of interpretations, two related problems arise with data sharing: 1) de-identification without decontextualization is often not sufficient and 2) decontextualization may lead to a misinterpretation of the data by secondary data analysts.

First, in personally and politically sensitive studies, the issue of confidentiality goes beyond the recognition of individual participants by their communities or social networks. Especially in cases of small and marginalized communities, such data can be intentionally (politically) misused to cause targeted harm. To think that the data are safe because the access is restricted to registered users (researchers) is irresponsibly naïve. Second, even without the premise of bad intentions, different levels of access do not solve the problem of sensitive data in the absence of mechanisms to enforce that the secondary data analysts are also, in addition to original depositors, held accountable to the research participants and the data. Without such mechanisms in place, sensitive qualitative data must sometimes be de-identified to such a degree that the data are actually decontextualized: key contextual information crucial for interpreting the data is removed. At the same time, qualitative analysts must be informed by these contextual data in their interpretations. Even if they cannot present this information in publications in order to protect the participants, it remains a part of the analytic process, and the analyst must explain how this information underlined their interpretations even if they cannot provide the specifics. However, if the data must be decontextualized to be shared, then qualitative interpretation is not possible for the secondary data analysts since the decontextualized interpretation would challenge the integrity of the process.

In summary, sharing qualitative data without decontextualization may be ethically and morally problematic. However, interpreting decontextualized data is epistemologically problematic. It also betrays the relationship of trust with the research participants who, even after consenting to sharing of the data, did not release the researchers from the responsibility entailed in that trust.

Conclusion

What are the lessons here for academic publishing? First, the goal of open science, and of academic publishing that supports it, should be to promote accountability and transparency. Emphatically, transparency should neither be conflated with data sharing nor should it serve reproducibility. Qualitative research can be fully transparent and also ethically responsible without making all the data available and accessible; reproducibility is a meaningless concept in this context. Therefore, a truly inclusive OS should position transparency in a way that is compatible with different epistemological positions of qualitative, quantitative, and mixed-methods research, respecting and incorporating the specificities of each.

Social science methods are complex because societies are complex. Therefore, funding agencies and academic journals should refrain from imposing the 'one-solution-fits-all' approach on researchers. Instead, we should seek to recognize good publishing practices that work towards recognizing the challenges of different types of data and analyses and finding different ways to accommodate transparency and accountability. Such publishing practices, for example, could be highlighted by professional associations and recommended for emulation, and they could also be specifically appreciated by funding agencies when guiding evaluators on how to judge the publication outlets.

Finally, we must be careful not to impose blindly and uniformly even such nice-sounding principles as FAIR on all types of studies. We must not sanction those who cannot or must not comply with the FAIR principles by limiting their options to do funded research and to publish. Working in our individual roles – as evaluators of funding applications, as peer reviewers of scholarly work, and as journal editors – it must become our duty and responsibility to demand that these issues be discussed and challenged at the institutional level until OS practices are expanded to truly serve the OS principles, instead of only one research epistemology.

References

1. European Commission. The EU's open science policy. https://ec.europa.eu/info/research-and-innovation/strategy/strategy-2020-2024/our-digital-future/open-science_en (accessed 12 January 2022).
2. Bezzak S, Clyburne-Sherin A, Conzett P, et al. The Open Science Training Handbook (1.0). FOSTER Open Science; 2018. <https://book.fosteropenscience.eu/> (accessed 12 January 2022).
3. Wilkinson MD, Dumontier M, Aalbersberg IJ, et al. The FAIR guiding principles for scientific data management and stewardship. *Scientific Data*. 2016;3(1):1-9.
4. Frias-Navarro D, Pascual-Llobell J, Pascual-Soler M, Perezgonzalez J, Berrios-Riquelme J. Replication crisis or an opportunity to improve scientific production? *Eur J Educ*. 2020;55(4):618-631.
5. Wasserman R. Ethical issues and guidelines for conducting data analysis in psychological research. *Ethics & Behavior*. 2013;23(1):3-15.
6. Rodgers BL, Cowles KV. The qualitative research audit trail: A complex collection of documentation. *Research in Nursing & Health*. 1993;16(3):219-226.
7. Bennett EA. Open science from a qualitative, feminist perspective: Epistemological dogmas and a call for critical examination. *Psychol Women Q*. 2021;45(4):448-456.

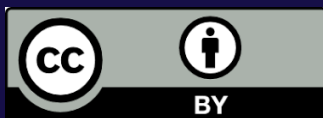
8. European Commission. Horizon Europe Programme Guide (version 1.2). European Commission. https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/guidance/programme-guide_horizon_en.pdf (accessed 3 November 2021).
9. Nosek BA, Alter G, Banks GC, et al. Guidelines for Transparency and Openness Promotion (TOP) in Journal Policies and Practices (Version 1.0.1). OSF. <https://osf.io/ud578/> (accessed 3 November 2021).
10. Center for Open Science. Open Science Badges enhance openness, a core value of scientific practice. <https://www.cos.io/initiatives/badges> (accessed 8 November 2021).
11. Schrag ZM. The case against ethics review in the social sciences. *Research Ethics*. 2011;7(4):120-131.
12. Tolich M. Are qualitative research ethics unique? In: Tolich M (ed). *Qualitative Ethics in Practice*. Routledge, 2016:33-47.
13. Tolich M. The making (s) of a qualitative code of ethics: Canada's tri-council policy statement: Ethical conduct for research involving humans. In: Tolich M (ed). *Qualitative Ethics in Practice*. Routledge, 2016:59-68.
14. Tolich M. Contemporary ethical dilemmas in qualitative research. In: Tolich M (ed). *Qualitative Ethics in Practice*. Routledge, 2016:25-32.
15. Ells C. Communicating qualitative research study designs to research ethics review boards. *Qualitative Report*. 2011;16(3):881-891.
16. Qualitative Data Repository. De-Identification. Syracuse University. <https://qdr.syr.edu/guidance/human-participants/deidentification> (accessed 4 November 2021).
17. Qualitative Data Repository. Access Controls. Syracuse University. <https://qdr.syr.edu/guidance/human-participants/access-controls> (accessed 4 November 2021).

ease / publications

ese / European Science Editing

European Science Editing is an official publication of EASE. It is an open access peer-reviewed journal that publishes original research, review and commentary on all aspects of scientific, scholarly editing and publishing.

<https://ese.arphahub.com/>
<https://www.ease.org.uk>
https://twitter.com/Eur_Sci_Ed
<https://www.linkedin.com/company/easeeditors/>



© 2022 the authors. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.